

**THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re the Application of : **Kanta YAMAMOTO**

Filed : **Concurrently herewith**

For : **PACKET TRANSMITTING APPARATUS**

Serial No. : **Concurrently herewith**

May 24, 2001

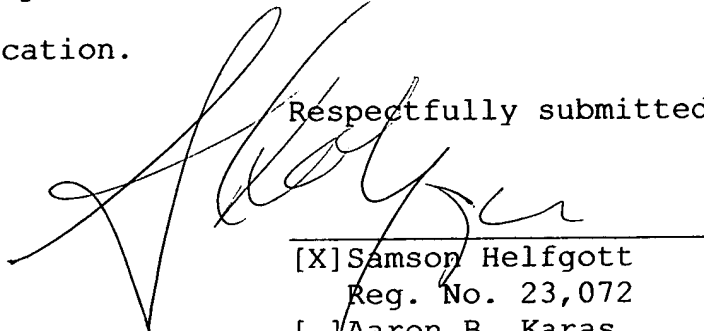
Assistant Commissioner of Patents  
Washington, D.C. 20231

**SUBMISSION OF PRIORITY DOCUMENT**

S I R:

Attached herewith is Japanese Patent Application No.  
2000-015615 of January 24, 2001 whose priority has been claimed  
in the present application.

Respectfully submitted



☒ Samson Helfgott  
Reg. No. 23,072  
☐ Aaron B. Karas  
Reg. No. 18,923

HELFGOTT & KARAS, P.C.  
60th FLOOR  
EMPIRE STATE BUILDING  
NEW YORK, NY 10118  
DOCKET NO.: FUSA 18.695  
BHU:priority

Filed Via Express Mail  
Rec. No.: EL639744054US  
On: May 24, 2001  
By: Brendy Lynn Belony  
Any fee due as a result of this paper,  
not covered by an enclosed check may be  
charged on Deposit Acct. No. 08-1634.



日 本 国 特 許 庁

PATENT OFFICE  
JAPANESE GOVERNMENT

J1046 U.S. PTO  
09/864995  
05/24/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application:

2001年 1月24日

出 願 番 号  
Application Number:

特願2001-015615

出 願 人  
Applicant (s):

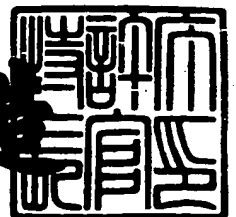
富士通株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 3月23日

特許庁長官  
Commissioner,  
Patent Office

及 川 耕 造



出証番号 出証特2001-3021987

【書類名】 特許願

【整理番号】 0000990

【提出日】 平成13年 1月24日

【あて先】 特許庁長官殿

【国際特許分類】 H04L 12/56

【発明の名称】 パケット送出装置

【請求項の数】 5

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 山本 幹太

【特許出願人】

【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【代理人】

【識別番号】 100084711

【弁理士】

【氏名又は名称】 齋藤 千幹

【電話番号】 043-271-8176

【手数料の表示】

【予納台帳番号】 015222

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704946

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 パケット送出装置

【特許請求の範囲】

【請求項 1】 優先度がそれぞれ異なる複数のグループに属するパケットを順に送出するパケット送出装置において、

グループ毎に待ち行列を生成し、各待ち行列を構成する要素に順番にパケット送出権を与える待ち行列制御部、

少なくとも 1 つの待ち行列構成要素に応じたパケットが送出待ちになっているグループのうち、優先度が最大のグループをパケット送出グループとして決定するパケット送出グループ決定部、

を備え、待ち行列制御部は前記決定されたパケット送出グループの待ち行列においてパケット送出権のある行列構成要素に応じたパケットを送出することを特徴とするパケット送出装置。

【請求項 2】 入力ポートと該ポートより入力するパケットに付加する各品質クラスの組み合わせを待ち行列構成要素とし、それぞれの待ち行列構成要素が属するグループを設定するグループ設定部、

を備え、前記待ち行列制御部は、

待ち行列毎に各待ち行列構成要素にラウンドロビン形式で順番に、かつ、均等にパケット送出権を与える、

ことを特徴とする請求項 1 記載のパケット送出装置。

【請求項 3】 待ち行列構成要素毎に送出待ちのパケットを記憶するバッファ部、

送出待ちパケットが記憶されているバッファ部に応じた待ち行列構成要素毎に送出要求のリクエスト信号を発生するリクエスト発生部、

を備え、前記送出グループ決定部は、各グループに属する待ち行列構成要素からのリクエスト信号の有無に基づいて送出待ちパケットが存在するグループを識別し、これらグループのうち優先度が最大のグループをパケット送出グループとして決定する、

ことを特徴とする請求項 1 記載のパケット送出装置。

【請求項 4】 行列構成要素毎に 1 回のパケット送出権で保証するデータ送出量を設定する保証データ量設定部、

パケット送出権を与えられた行列構成要素に応じたパケットの実際の送出量を監視する監視部、

実際のデータ送出量が前記保証データ量と等しくなったとき、送出権を次の行列構成要素に渡すための制御信号を出力する制御信号発生部、

を備え、前記待ち行列制御部は、該制御信号に基づいてパケット送出権を次の行列構成要素に付与する、

ことを特徴とする請求項 1 または 2 記載のパケット送出装置。

【請求項 5】 待ち行列構成要素毎に設定時間当りに送出するパケットのデータ送出流量を設定するデータ送出流量設定部、

待ち行列構成要素毎に前記設定時間当りのパケットの実際のデータ送出流量を監視する手段、

待ち行列構成要素毎にパケットの実際のデータ送出流量を監視し、該データ送出流量が前記設定データ送出流量に等しくなった時、前記設定時間が経過するまで該待ち行列構成要素に応じたパケットの送出を禁止する送出禁止信号を発生するパケット送出禁止手段、

を備えたことを特徴とする請求項 1 又は 2 記載のパケット送出装置。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は優先度がそれぞれ異なる複数のグループに属するパケットを順に送出するパケット送出装置に係わり、特に、可変長パケットのスイッチングを行う IP ルータ装置、 Ethernet スイッチ装置などのパケット送出装置に関する。

【0 0 0 2】

【従来の技術】

IP ネットワーク (インターネット) が登場した頃、ネットワーク内では差別化処理などではなく、ある特定のフローに帯域を優先して割り当てるなどの処理は無かった。この初期のインターネット形態では、帯域の割り当ては「早い者勝ち」の世

界であり、多くデータを送れば送るほどネットワークの帯域を占有するというもので、スケジューリング論理も単純なFiFo(First In First Out)であった。

しかし、時代は次第にIPネットワークに、電子ファイル・データ、e-mail、静止画像だけではなく、動画、音声というマルチメディア情報を送れる事を期待しはじめた。動画、音声というマルチメディア情報は、いわゆる”リアルタイム・トラフィック”と呼ばれるように、送信元から送信先に対して、「ある一定帯域、一定周期で、データの損失無し」で、通信される事が期待されているトラフィックである。なお、これまでのリアルタイム・トラフィックは、送信元から送信先への固定的な帯域を確保してから通信する、いわゆる電話網にてやり取りされていたトラフィックであった。

#### 【0003】

また、急激に増加したインターネット・ユーザからの膨大なパケット・データは、ネットワーク内でパケットの混雑状態を招き、多くのパケットがネットワーク内で処理できず廃棄され始めた。そうすると、多くのユーザにおいて、目に見て分かるくらいのthroughputの低下が感じられるようになった。ここに、「私は人より多くお金を払うので、throughputを上げてくれ」という要望が生まれる。これはよく、車の一般道と高速道路に例えられる。固定的な帯域を保証してもらうのは高速道路、いろんなトラフィックが帯域の奪い合いをやっているのが一般道、という事だが、注意が必要なのは、IPネットワークの場合、高速道路のトラフィックを保証すればするだけ、ネットワークとしてのリソース(帯域)を一般道から奪っているの、一般道はさらに混在する、という点である。

#### 【0004】

上記の背景は、もともと「統計多重による帯域の有効活用」を旨とするIPネットワークに、帯域制御、優先度制御の必要性を生んだ。インターネットにおける差別化思想は、QoS(Quality of Service)、Diffserve、Intserveなどにより、サービス・ポリシーとして定義、確立されようとしている。差別化思想、サービスポリシーでは、パケット・フロー識別規定(classification)、入力トラフィック特性規定、パケット廃棄優先度規定、帯域保証規定などにより、各々のパケット・フローに対するサービスを管理する。

帯域制御方式の代表的な例として、ストリクト・キューイング(Strict Queuing)と、重み付けキューイング(WFQ:Weighted Fair Queuing)がある。

ストリクト・キューイングは、優先度が高く設定されたパケットがあれば、必ず優先的に先に送出する単純優先度方式である。図 8 は単純優先度方式の説明図であり、 $1_1 \sim 1_4$  はクライアント、2 はルータ、3 はアプリケーションサーバであり、各クライアント  $1_1 \sim 1_4$  がルータ 2 を介してサーバ 3 にパケットを送信する場合であり、P は優先度を表し、値が大きいほど優先度が高くなっている。ルータ 2 は同時に各クライアント  $1_1 \sim 1_4$  からパケット①～④が各ポートに入力すると優先度が高い順に、図 3 では③→②→④→①の順にサーバ 3 に送出する。

#### 【 0 0 0 5 】

図 8 は単純な場合であるが、一般的に優先度制御を行うと、純粋なFiFo(First In First Out)の順序でパケットは送出されなくなる。図 9 では、①のパケットが⑤のパケットより先に到着しているにも関わらず、⑤が先に送出されている事がわかる。以上のように、ストリクト・キューイングでは、優先度の高いパケットが連続して到着した場合、優先度の低いパケットは、優先度の高いパケットが送出し終わるまで、送出されることはない。このため、優先度の低いパケットはキューイング・メモリにて滞留する可能性が高く、スレッシュホルドを超えて入力があった場合パケット廃棄となる。

ストリクト・キューイング方式の特徴としては、非常に厳格な優先度制御が行われ、スケジューリング処理も簡単、単純であるが、各フロー間に自由度のある帯域設定、優先度設定を行うことができない。

#### 【 0 0 0 6 】

重み付けキューイング方式は、優先度順位をアナログ的なパラメータにて、優先比率(重み付け: weight)を決め、決められた比率に従ってパケットを送出する方式である。図 1 0 及び図 1 1 は重み付けキューイング方式の説明図であり、サーバ 3 への一つの出力ポートに対して、クライアント  $1_1 = 12.5\%$ 、クライアント  $1_2 = 12.5\%$ 、クライアント  $1_3 = 25\%$ 、クライアント  $1_4 = 50\%$ 、という優先比率が割り当てられている。重み付けキューイング方式では、ストリクト・キューイングのように特定の優先度に厳格に偏ることはなく、この比率に従った帯域割り当て

処理が行われる（図 1 1 参照）。

しかし、重み付けキューイング方式は可変長パケットにおける優先比率制御であり、パケット転送数で重み付け制御すると、最小パケットと最大パケットのサイズの違いが認識されない。Ethernetフレームを例にとると最小64バイト、最大1522バイトなので、ルータ2において図12に示すようにクライアント1<sub>1</sub>、1<sub>2</sub>からのパケットを交互にサーバ3に送出するものとする、優先比率の誤差が最大24倍近い開きが生じてしまう。この解決方法としては、図13に示すように可変長パケットを固定長サイズに分割して、固定長データの転送数にて制御する方法がある。

#### 【0007】

しかし、この方式では、ルータ2の図示しないルーチング部あるいはスイッチ部の前段にsegment処理部（分割処理部）、後段にre-assembly処理部（組立て処理部）が必要となる問題がある。特に、後段での組立て処理部において同時に複数のパケットが組み立て中となる事を想定しなければならず、その各々のパケット毎に論理的、もしくは物理的に個別な組立てバッファを用意し、制御する必要がある問題がある。図14はルータのパケット送出制御説明図であり、2aはパケットスイッチ部、2bはスケジューラ、2c<sub>1</sub>～2c<sub>4</sub>は各クライアントからのパケットを組み立てる組立てバッファ、2dは各バッファからパケットを読み出して送出するパケット読み出部である。

#### 【0008】

##### 【発明が解決しようとする課題】

##### (1) 第1の課題

重み付けキューイング (Weighted Fair Queueing)方式はパケット単位に Virtual Clockという概念を導入する。このVirtual Clockとは、パケットが装置から出力されるまでの時間、すなわち、装置内にパケットが存在する時間である。Virtual Clock情報は、装置内の個々のパケット毎に個別に管理しなければならず、その生成、管理は膨大な処理となる。換言すれば、重み付けキューイング方式におけるスケジューラは、装置内部に存在する全パケットに対して大量かつ複雑な制御情報を生成、管理しなければならず、スケジューリング処理が複雑に



なり、しかも処理時間が長くなる問題がある。

【0 0 0 9】

(2) 第2の課題

可変長パケットをATMセルのような固定長データに分割し、スケジューラも固定長データ単位にスケジューラ管理を行うように作り込むと、スケジューラの制御は比較的簡単にはなる。しかし、そのためには、ルーチング部、スイッチ部の前段に可変長パケットの分割処理が必要となり、かつスイッチ部の後段では組み立て処理が必要となる。そうすると、後段での可変長パケット・組み立てバッファでは、同時に複数のパケットが組み立て中になる事を想定しなければならず、その各々のパケット毎に論理的、もしくは物理的に個別な組み立てバッファを用意する必要があり、装置規模が大きくなる問題がある。

【0 0 1 0】

(3) 第3の課題

物理帯域を100%有効に使用できることが望ましい。このため、出力先に可変長パケットを無駄な空き帯域無しに、しかも時間的に隙間なくパケット・データを詰めて出力する必要がある。しかし、従来のように複雑なスケジューラ構成では、1回のスケジューリング処理時間が長くなり、出力先へ隙間なくパケット・データを出力することができない。

【0 0 1 1】

(4) 第4の課題

従来のスケジューラは、スケジューリング対象となる待ち行列 (Queue) を1段階で処理しており、優先度設定も全ての行列構成要素 (例えばパケット) を対象として設定されている。この従来方式では、転送許可 (grant) を各行列構成要素間にて巡回させるためにRound-Robin方式を採用している。Round-Robin方式は、ある待ち行列構成要素が転送許可 (grant) を獲得したら、次のスケジューリング処理において該待ち行列構成要素の優先度を最も低くする方式である。Round-Robin方式は公正 (fair) なスケジューリングには向いているが、非線形な帯域割り当て制御 (つまり、差別化サービス) にはそのままでは採用できない。例えば、帯域を高め設定されている行列構成要素に対しては、転送許可の巡回頻度を

上げる必要があるが従来のRound-Robin方式のスケジューラはかかるスケジューリング制御ができない。

【 0 0 1 2 】

(5) 第 5 の課題

可変長パケットを取り扱う物理帯域保証型スケジューラは、大きく 2 種類に分類される。一つは可変長パケット・データを固定長データに分割して行う方法、もう一つは可変長パケット・データのまま行う方法である。本発明では、可変長パケット・データのままで行う方法を採用する。かかる方法において、スケジューラからの転送許可(grant)は、パケット単位で出力されるのが一般的である。この場合、1 回の転送許可において転送されるパケット・データ長は一定ではない。つまり、各行列構成要素において等しく 1 回の転送許可を巡回したとしても、同じだけの物理帯域を占有したことにはならない。これが数十回、数百回となると、最悪同じ回数だけ転送許可を巡回させているにも関わらず、行列構成要素への割り当て帯域に不公平(unfair)が生じていしまう。これが、可変長パケットを扱うスケジューラにおいて、帯域保証制御を難しくしている原因である。

【 0 0 1 3 】

(6) 第 6 の課題

上記の第 5 の課題は、1 回の転送許可における伝送保証帯域を規定する事により、可変長パケット伝送における不公平感を無くせる。この場合、伝送保証帯域に達する前に送出すべきパケットが存在しなくなった場合の処理を規定する必要がある。かかる処理方式としては、パケットが来るまで待つ方式がある。これは、伝送保証帯域に達するまで、対象行列構成要素に連続して転送許可(grant)を与え続ける方式であるが、物理帯域の 100% 利用は不可能であり、無駄な空き帯域が生じる問題がある。

【 0 0 1 4 】

(7) 第 7 の課題

前述しているとおり、可変長パケット・スケジューラは、公平な帯域割り当て制御を実現するためにパケット転送個数情報を使用することはできない。なぜならば、同じ 10 個のパケット・データでも一方の行列構成要素からは最小サイズの

パケットが多く、もう一方の行列構成要素からは最大サイズのパケットが多い場合、同じ帯域設定としているのに不公平(unfair)が生じるからである。従って、可変長パケットであっても厳密な帯域制御が行えるようにする必要がある。

【 0 0 1 5 】

以上より、本発明の目的はストリクトキューイング方式及び重み付けキューイング方式を併用して可変長パケットを分割することなく帯域制御を可能にし、スケジューラ回路規模の縮小化、高速処理化を図ることである。

本発明の別の目的は、入力ポートと該ポートより入力するパケットに付加される品質クラスとの組み合わせを待ち行列構成要素とし、パケット単位でなくこの待ち行列構成要素単位で帯域制御、優先度制御することによりスケジューラ回路規模の縮小化、高速処理化を図ることである。

本発明の別の目的は、ルーチング部やスイッチ前段のパケット分割処理や後段のパケット組立て処理を不要にすることである。

本発明の別の目的は、1回のスケジューリング処理時間が短くなるように、すなわち、スケジューリング処理時間が最小パケットの伝送時間以内で収まるようにすることである。

【 0 0 1 6 】

本発明の別の目的は、待ち行列構成要素を数段階の絶対優先度グループにグルーピングし、グループ内でRound-Robin方式により各待ち行列構成要素に順番にパケット送出権を与え、これにより、単純なRound-Robin方式であっても、優先度の高い特定の行列要素に対して転送許可の巡回頻度を上げれるようにすることである。

本発明の別の目的は、優先度グループを単純な絶対優先度方式により選択することにより、シンプルなハードウェアで高速スケジューリング処理を可能にすることである。

本発明の別の目的は、各優先度グループに属する行列構成要素を任意に設定でき、また、その数に制限を設けないようにでき、しかも、優先度の高いグループの流量と、優先度の低いグループの流量の比率が偏っていても、帯域制御を可能にすることである。

## 【 0 0 1 7 】

本発明の別の目的は、各行列構成要素に対する 1 回の転送許可による伝送保証帯域を設定して各行列構成要素に割り当てる帯域の制御を行えるようにすることである。

本発明の別の目的は、連続して送出するパケットが存在しなくなれば、直ちにその行列構成要素に対する転送許可を解除し、他の行列構成要素に転送許可を渡し、共有している物理帯域において無駄な空き帯域をなくし、帯域を 100% 利用できるようにすることである。

本発明の別の目的は、各入力ポートの品質クラス（行列構成要素）毎に所定時間当りに送出するパケットのデータ送出流量（帯域）を設定することにより厳密な帯域制御を実現することである。

## 【 0 0 1 8 】

## 【課題を解決するための手段】

本発明の第 1 のパケット送出装置は、(1) 優先度がそれぞれ異なる複数のグループの各グループ毎に待ち行列を生成し、各待ち行列を構成する要素に順番にパケット送出権を与える待ち行列制御部、(2) 少なくとも 1 つの待ち行列構成要素に応じたパケットが送出待ちになっているグループのうち、優先度が最大のグループをパケット送出グループとして決定するパケット送出グループ決定部、を備え、待ち行列制御部は該パケット送出グループの待ち行列においてパケット送出権のある行列構成要素に応じたパケットを送出する。

## 【 0 0 1 9 】

本発明の第 2 のパケット送出装置は、更に、(3) 入力ポートと該ポートより入力するパケットに付加する各品質クラスの組み合わせを待ち行列構成要素とし、それぞれの待ち行列構成要素が属するグループを設定するグループ設定部を備え、前記待ち行列制御部は、待ち行列毎に各待ち行列構成要素にラウンドロビン形式で順番に、かつ、均等にパケット送出権を与える。

## 【 0 0 2 0 】

本発明の第 3 のパケット送出装置は、更に、(4) 待ち行列構成要素毎に送出待ちのパケットを記憶するバッファ部、(5) 送出待ちパケットが記憶されているパ

ツファ部に応じた待ち行列構成要素毎に送出要求のリクエスト信号を発生するリクエスト発生部を備え、前記送出グループ決定部は、各グループに属する待ち行列構成要素からのリクエスト信号の有無に基づいて送出待ちパッケージが存在するグループを識別し、これらグループのうち優先度が最大のグループをパッケージ送出グループとして決定する。

#### 【 0 0 2 1 】

本発明の第4のパッケージ送出装置は、更に、(6) 行列構成要素毎に1回のパッケージ送出権で保証するデータ送出量を設定する保証データ量設定部、(7) パッケージ送出権を与えられた行列構成要素に応じたパッケージの実際の送出量を監視する監視部、(8) 実際のデータ送出量が前記保証データ量と等しくなったとき、送出権を次の行列構成要素に渡すための制御信号を出力する制御信号発生部を備え、前記待ち行列制御部は、該制御信号に基づいてパッケージ送出権を次の行列構成要素に付与する。

#### 【 0 0 2 2 】

本発明の第5のパッケージ送出装置は、更に、(9) 待ち行列構成要素毎に設定時間当りに送出するパッケージのデータ送出流量を設定するデータ送出流量設定部、(10) 待ち行列構成要素毎に前記設定時間当りのパッケージの実際のデータ送出流量を監視する手段、(11) 待ち行列構成要素毎にパッケージの実際のデータ送出流量を監視し、該データ送出流量が前記設定データ送出流量に等しくなった時、前記設定時間が経過するまで該待ち行列構成要素に応じたパッケージの送出を禁止する送出禁止信号を発生するパッケージ送出禁止手段を備えている。

#### 【 0 0 2 3 】

上記本発明のパッケージ送出装置は、どの優先度グループに転送許可(grant)を出すかという処理と、その優先度グループに属するどの行列構成要素に転送許可(grant)を出力するかという処理との2段構成でパッケージを送出するため、パッケージ送出制御を簡単な構成で行うことができる。

又、本発明のパッケージ送出装置は、可変長パッケージデータを固定長に分割することなく、可変長のままでスケジューリング処理を行い、各行列構成要素毎に、すなわち、各入力ポートの品質クラス毎に正確かつ、自由に帯域割り当て制御を

実現できる。

【 0 0 2 4 】

又、本発明の packets 送出装置は、packets 単位でなく、各入力ポートの品質クラス単位でスケジューラ部へ転送要求信号 (request) を生成することにより、大幅な回路規模の縮小化、高速化が可能である。

又、本発明の packets 送出装置は、固定長に分割されたデータ単位に対して転送許可 (grant) するのではなく、あくまで可変長 packets・データ単位に転送許可するため、packets・データ自体を分割すること無く転送でき、可変長 packets の分割処理 (segment 処理)、組立て処理 (re-assembly 処理) を必要としない。

又、本発明の packets 送出装置は、スケジューラをハードウェアで構成しているため、高速のスケジューリング判定を可能とし、1 回のスケジューラ判定時間を最小 packets 長の物理的な伝送時間以内に押えることが可能となる。これにより、スケジューラ処理によるディレイを少なくし、最小 packets・サイズが連続して入力された場合でも、無駄な物理的空き帯域を無くして連続的に packets の送出が可能である。

【 0 0 2 5 】

又、本発明の packets 送出装置は、優先度グループを自由に設定できるようにしたから、優先度の高い行列構成要素数と低い優先度の行列構成要素数の比率が偏っていても対応可能である。

又、本発明の packets 送出装置は、転送許可された行列構成要素に応じた packets は伝送保証帯域を超えるまで転送許可されるため、packet-by-packet 単位の装置であるにも関わらず、一回の転送許可において物理的な帯域保証 (byte/sec、もしくは bit/sec) が可能である。又、実際の帯域が伝送保証帯域となる前にすべての packets の送出が完了すれば直ちに転送許可を解除することにより、無駄な物理的空き帯域が発生するのを防止できる。

又、本発明の packets 送出装置は、基本的に単純優先度制御の Round-Robin スケジューラであるが、packets 毎の流量情報をフィードバックする構成を採用することにより、ハードウェアによる高速スケジューリング処理が可能となり、しかも、byte/sec、もしくは bit/sec という単位で絶対物理帯域の割り当て制御が可

能となる。

# 【0026】

## 【発明の実施の形態】

### (A) 第1実施例

#### (a) パケット送出装置の全体の構成

図1は本発明のパケット送出装置の構成図である。このパケット送出装置は、パケットを伝送路に送出する際、該パケットが到来した入力ポート $P_0 \sim P_n$ とその品質クラスとに基づいて所定のスケジューリングに従って送出するものである。

図1において、10はn個の入力ポートと1個の出力ポート有するパケット送出装置、11<sub>0</sub>～11<sub>n</sub>は入力伝送路、12は出力伝送路、 $P_0 \sim P_n$ は入力ポートである。パケット送出装置10は入力ポート $P_0 \sim P_n$ と該入力ポートから入力するパケットの品質クラス（サービスクラス）との組み合わせに基づいて優先制御及び帯域制御を行いながらパケットを出力伝送路に送出する。

# 【0027】

パケット送出装置10は、SONETの場合であれば各伝送路から入力する伝送路信号を光電変換すると共に、SONET終端処理、パケット分離処理などのインタフェース制御を行う入力側IF部21<sub>0</sub>～21<sub>n</sub>、入力側IF部から入力するパケットに装置内ヘッダの付加などの処理を施して出力するフォワーディング部22<sub>0</sub>～22<sub>n</sub>、フォワーディング部から入力するパケットをQoS毎に振り分けてキューイングするingress側キューイング部23<sub>1</sub>～23<sub>n</sub>、スケジューリング制御に基づいてキューイング部23<sub>0</sub>～23<sub>n</sub>の各キューイングバッファに記憶されているパケットを送出するスケジューリング部24、スケジューリング部から出力するパケットをキューイングして順番に出力するegress側キューイング部25、電光変換、装置内ヘッダの削除、SONETフレーム信号の生成等を行って出力伝送路に送出する出力IF部26を備えている。

# 【0028】

フォワーディング部22<sub>0</sub>～22<sub>n</sub>は図2に示すようにIPパケットPKTに装置内ヘッダFIH(Frame Information Header)を付加して出力するもので、装置内ヘッダFIHは、たとえば、①ルーティングフラグ/マルチキャストフラグ/廃棄フラグ、

②受信ポート番号、③出力ポート番号、④QoS(Quority of Service)、⑤パケット長、⑥マルチキャストの出力ポート等を含んでいる。このうちQoSには例えば以下の4つのサービスクラス(品質クラス)、

- ①iGS(Guranteed Serviceクラス)、
- ②iCLS(Controlled Load Serviceクラス)、
- ③iBE(Best Effortクラス)、
- ④iSpare(予備クラス)、

がある。iGSは要求された帯域を保証するサービスクラス、iCLSは所定の最低帯域を保証するサービスクラス、iBEは品質規定がないサービスクラス、iSpareは予備クラスである。各フォワーディング部 $22_0 \sim 22_n$ は、IPアドレス(発信元アドレス/宛先アドレス)、受信ポート番号、レイヤ4プロトコル(TCP/UDP)、サービス種別(WEB/FTP/TELNET)のいずれか、あるいはこれらの組み合わせに基づいてサービスクラスを決定してパケットに付加する。

#### 【0029】

ingress側キューイング部 $23_0 \sim 23_n$ はそれぞれサービスクラス毎に総計4つのキューイングバッファ $QBF_{i1} \sim QBF_{i4}$  ( $i=0 \sim n$ )を備え、対応するフォワーディング部 $22_0 \sim 22_n$ から入力するパケットのQoSを参照し、サービスクラス毎にパケットをキューイングバッファ $QBF_{i1} \sim QBF_{i4}$ に記憶する。又、ingress側キューイング部は、スケジューリング部24からの転送許可により順番にパケットを送出する。

#### 【0030】

##### (b) スケジューリング部

スケジューリング部24はキューイング部 $23_0 \sim 23_n$ の各キューイングバッファ $QBF_{01} \sim QBF_{04} \dots QBF_{71} \sim QBF_{74}$ にキューイングされているパケットをスケジューリング制御により順番に送出するもので、図3、図4に示す構成を備えている。尚、図3にはingress側キューイング部のキューイングバッファ $QBF_{01} \sim QBF_{04} \dots QBF_{71} \sim QBF_{74}$ を含めて示している。

図3、図4は図1におけるスケジューリング部24の構成図であり、二分した一方を図3に、他方を図4に示している。スケジューリング部24は大別する



と、ポート対応部 3 1<sub>0</sub> ~ 3 1<sub>7</sub> (n=7)、スケジューラ 3 2、パケット合成部 3 3 を有し、各ポート対応部 3 1<sub>0</sub> ~ 3 1<sub>7</sub> は同一の構成を備えている。

#### 【 0 0 3 1 】

##### (b-1) ポート対応部

ポート対応部 3 1<sub>0</sub> は、ポート P<sub>0</sub> から入力するパケットをサービスクラス（品質クラス）毎に振り分けてキューイングし、スケジューラ 3 2 からの転送許可に基づいてパケットを送出する部分で、GS用バッファ QBF<sub>01</sub>、CLS用バッファ QBF<sub>02</sub>、BE用バッファ QBF<sub>03</sub>、Spare用バッファ QBF<sub>04</sub> を備えている。同様に、ポート対応部 3 1<sub>i</sub> (i=1~7) はポート P<sub>i</sub> (i=1~7) から入力するパケットをサービスクラス毎に振り分けてキューイングし、スケジューラ 3 2 からの転送許可に基づいてパケットを送出する部分であり、GS用バッファ QBF<sub>i1</sub>、CLS用バッファ QBF<sub>i2</sub>、BE用バッファ QBF<sub>i3</sub>、Spare用バッファ QBF<sub>i4</sub> を備えている。

各ポート対応部 3 1<sub>i</sub> (i=0~7) は上記構成に加えて、リクエスト信号発生部 3 1 a、リクエスト禁止部 3 1 b、転送許可部（Grant部）3 1 c、連続転送許可部 3 1 d、伝送保証帯域記憶部 3 1 e<sub>1</sub> ~ 3 1 e<sub>4</sub>、レジスタ 3 1 f、設定値ロード部 3 1 g、オクテットカウンタ 3 1 h、比較部 3 1 i、合成部 3 1 j を有している。

#### 【 0 0 3 2 】

リクエスト信号発生部 3 1 a は、各キューイングバッファ QBF<sub>01</sub> ~ QBF<sub>04</sub> にそれぞれ転送待ちパケットが存在するとき、転送許可を要求するリクエスト信号 (Request) R<sub>11</sub> ~ R<sub>14</sub> を発生する。リクエスト禁止部 3 1 b は、所定サービスクラスのパケットを設定時間内に設定量以上送出したとき、該サービスクラスのパケットの送出を禁止するためにリクエスト信号をマスクし、リクエスト信号がスケジューラ 3 2 に入力するのを禁止する。転送許可部（Grant部）3 1 c はスケジューラ 3 2 から入力する転送許可信号 (Grant) に基づいてキューイングバッファ QBF<sub>01</sub> ~ QBF<sub>04</sub> にパケットの転送を許可する。連続転送許可部 3 1 d は連続してパケットの転送を許可信号を送出すると共に、所定の条件が成立したとき次の行列構成要素にパケット送出権を与えるようスケジューラ 3 2 に巡回信号 (Rotate) を入力する。所定の条件とは、① 1 回の転送許可により送出可能なデータ量 (伝送

保証データ量)を送出したとき、②実際の送出データ量が1回の転送許可により送出可能な伝送保証データ量に等しくなる前にパケットがバッファからすべて送出されたときであり、いずれかの条件が成立すると、転送許可信号(Grant)の出力を停止すると共に、巡回信号(Rotate)を出力する。

#### 【0033】

伝送保証データ量記憶部31e<sub>1</sub>~31e<sub>4</sub>は1回の転送許可により送出可能な伝送保証データ量(byte/sec)TH<sub>1</sub>~TH<sub>4</sub>をサービスクラス毎に記憶し、設定値ロード部31gは転送許可された入力ポートのサービスクラスに応じた設定データ量TH<sub>i</sub>をレジスタ31fにロードする。オクテットカウンタ31hは実際に送出されるパケットデータのオクテット数をカウントし、比較部31iは実際に送出されるパケットのオクテット数とレジスタ31fに設定された伝送保証データ量TH<sub>i</sub>を比較し、一致したときStop信号を出力すると共に、オクテットカウンタ31hをリセットする。合成部31jは各キューイングバッファQBF<sub>01</sub>~QBF<sub>04</sub>から出力するパケットを合成してパケット合成部33を介してegress側キューイング部25に送出する。

#### 【0034】

##### (b-2) スケジューラ

スケジューラ32において、レジスタ32a<sub>01</sub>~32a<sub>74</sub>には入力ポートP<sub>0</sub>~P<sub>7</sub>のサービスクラス(iGS, iCLS, iBE, iSpare)毎に一定時間内に送出可能な最大データ量TH<sub>i1</sub>~TH<sub>i4</sub>(byte/sec)(i=0~7)が設定される。カウンタ32b<sub>01</sub>~

32b<sub>74</sub>は各入力ポートP<sub>0</sub>~P<sub>7</sub>のサービスクラス(iGS, iCLS, iBE, iSpare)毎に設定時間内に実際に送出したパケットデータ量をカウントするものである。フロー識別部32cはパケットに付加された装置内ヘッダFIHを参照して送出パケットの入力ポート番号及びサービスクラスを識別し、識別した入力ポート番号/サービスクラスに応じたカウンタ32b<sub>01</sub>~32b<sub>74</sub>により送出データ量をカウントさせる。比較部32e<sub>01</sub>~32e<sub>74</sub>は、各入力ポートの品質クラス毎のパケットデータ送出量が設定最大データ量TH<sub>i1</sub>~TH<sub>i4</sub>と等しくなった時、すなわち、カウンタ32b<sub>01</sub>~32b<sub>74</sub>の計数値が設定最大データ量TH<sub>i1</sub>~TH<sub>i4</sub>に等しくなったとき、設定時間が経過するまで該パケットの送出を禁止する送出禁止信号

INHを発生する。リセット部 3 2 f は設定時間毎に各カウンタの計数値をリセットする。

#### 【 0 0 3 5 】

リクエスト集積部 3 2 g はポート対応部 3 1<sub>0</sub> ~ 3 1<sub>7</sub> から出力する 3 2 個のリクエスト信号を集積して、順番に並べて出力し、グループ設定部 3 2 h は、第 1 ~ 第 4 優先グループを設定すると共に、入力ポートと該ポートより入力するパケットのサービスクラスとの組み合わせを待ち行列構成要素とし、これら各行列構成要素が属するグループを設定する。図 5 (A) の例では、

(1) 各入力ポート P<sub>0</sub> ~ P<sub>7</sub> におけるサービスクラス GS を第 1 優先度グループの行列構成要素 P<sub>0</sub>-GS ~ P<sub>7</sub>-GS とし、

(2) 各入力ポート P<sub>0</sub> ~ P<sub>7</sub> におけるサービスクラス CLS を第 2 優先度グループの行列構成要素 P<sub>0</sub>-CLS ~ P<sub>7</sub>-CLS とし、

(3) 各入力ポート P<sub>0</sub> ~ P<sub>7</sub> におけるサービスクラス BE を第 3 優先度グループの行列構成要素 P<sub>0</sub>-BE ~ P<sub>7</sub>-BE とし、

(4) 各入力ポート P<sub>0</sub> ~ P<sub>7</sub> におけるサービスクラス Spare を第 4 優先度グループの行列構成要素 P<sub>0</sub>-Spare ~ P<sub>7</sub>-Spare とする。

#### 【 0 0 3 6 】

図 5 (B) は、

(1) 各入力ポート P<sub>0</sub> ~ P<sub>6</sub> におけるサービスクラス GS を第 1 優先度グループの行列構成要素 P<sub>0</sub>-GS ~ P<sub>6</sub>-GS とし、

(2) 各入力ポート P<sub>0</sub> ~ P<sub>7</sub> におけるサービスクラス CLS、入力ポート P<sub>7</sub> のサービスクラス GS、入力ポート P<sub>0</sub> のサービスクラス BE をそれぞれ第 2 優先度グループの行列構成要素 P<sub>7</sub>-GS, P<sub>0</sub>-CLS ~ P<sub>7</sub>-CLS, P<sub>0</sub>-BE とし、

(3) 各入力ポート P<sub>1</sub> ~ P<sub>7</sub> におけるサービスクラス BE を第 3 優先度グループの行列構成要素 P<sub>1</sub>-BE ~ P<sub>7</sub>-BE とし、

(4) 各入力ポート P<sub>0</sub> ~ P<sub>7</sub> におけるサービスクラス Spare を第 4 優先度グループの行列構成要素 P<sub>0</sub>-Spare ~ P<sub>7</sub>-Spare とする。

以上のように、各優先度グループには適宜必要な行列構成要素を含ませることができる。

## 【 0 0 3 7 】

優先度マッピング部 3 2 i は 3 2 個のリクエスト信号を優先度グループ毎に仕分けして出力する。すなわち、優先度マッピング部 3 2 i は 3 2 個のリクエスト信号を、(1) 第 1 優先度グループに属する入力ポート/QoS(行列構成要素)に応じたリクエスト信号、(2) 第 2 優先度グループに属する行列構成要素に応じたリクエスト信号、(3) 第 3 優先度グループに属する行列構成要素に応じたリクエスト信号、(4) 第 4 優先度グループに属する行列構成要素に応じたリクエスト信号のそれぞれに分類して出力する。

絶対優先度判定部 3 2 j は絶対優先度判定処理を行い、所属する少なくとも 1 つの行列構成要素から転送要求のリクエスト信号(Request)が発生しているグループのうち優先度が最大のグループをパケット送出グループとして決定し、該グループにenable信号(有効信号を)を出力する。すなわち、第 1 優先グループに所属する行列構成要素からリクエスト信号が発生していれば、該第 1 優先グループにイネーブル信号を入力する。しかし、第 1 優先グループに所属する行列構成要素からリクエスト信号が発生せず、第 2 優先グループに所属する行列構成要素からリクエスト信号が発生していれば、該第 2 優先グループにイネーブル信号を入力する。

## 【 0 0 3 8 】

各優先グループ毎に設けられたラウンドロビン回路 3 3 k<sub>i</sub> (i=1~4) は第 i 優先度グループの各行列構成要素で待ち行列を生成し、各待ち行列構成要素にラウンドロビン方式で順番に転送許可(パケット送出権)を与える。Round-Robin方式は、基本的な巡回方向に順番に、かつ、均等に転送許可(パケット送出権)を行列構成要素に与え、かつ、前回転送許可した行列構成要素を今回では最低優先度となるようにするルールで転送許可を与える方式である。基本的な巡回方向は、優先度グループ設定時に固定的に設定される。ここで注意する事は、転送許可は、あくまで送出待ちパケットが存在する待ち行列構成要素(リクエスト信号を出力している待ち行列構成要素)に対して出力される事である。

## 【 0 0 3 9 】

図 6 はRound-Robin方式の説明図であり、図 5 (A) の第 1 優先グループの待

ち行列を示している。斜線で示す位置が行列先頭で、優先度が最も高く、以下順に時計方向に優先度が低くなっており、図 6 (A)では、行列構成要素 $P_0$ -GSの優先度が最高で、行列構成要素 $P_7$ -GSの優先度が最低になっている。

Round-Robin方式では、 $P_0$ -GS $\rightarrow P_1$ -GS $\rightarrow P_2$ -GS $\rightarrow P_3$ -GS $\rightarrow P_4$ -GS $\rightarrow P_5$ -GS $\rightarrow P_6$ -GS $\rightarrow P_7$ -GS $\rightarrow P_0$ -GS $\rightarrow \dots$ の順番に、かつ、均等に各行列構成要素に転送許可（パケット送出権）を与える。すなわち、第 1 優先グループが図 6 (A) に示す状態にあるとき、この第 1 優先グループに転送許可が出されると(enable="1"), 先頭の行列構成要素 $P_0$ -GSにパケット送出権を与え(Grant)、該行列構成要素 $P_0$ -GSに応じた送出待ちパケットを一定量送出させる。又、送出完了により巡回信号(Rotate)が発生すると、図 6 (B) に示すように、各行列構成要素の位置を反時計方向に 1 づつシフトし、次の $P_1$ -GSを先頭の行列構成要素とし、 $P_0$ -GSを最低優先度の行列構成要素とする。そして、再び第 1 優先グループに転送許可が出されると、先頭の行列構成要素 $P_1$ -GSにパケット送出権を与え、該行列構成要素 $P_1$ -GSに応じたパケットを一定量送出させる。以後、同様に順番に、かつ、均等に各行列構成要素に転送許可（送出権）を与える。ただし、先頭の第 1 優先度の行列構成要素に応じた送出待ちパケットが存在しなければ、すなわち、該行列構成要素から転送要求のリクエスト信号(Request)が発生していなければ、第 2 優先度の行列構成要素(図 6 (A) では $P_1$ -GS)にパケット送出権を与える。又、第 2 優先度の行列構成要素 $P_1$ -GSに応じた送出待ちパケットが存在しなければ、第 3 優先度の行列構成要素 $P_2$ -GSにパケット送出権を与え、以下同様に送出待ちパケットが存在する行列構成要素にパケット送出権を与える。

【 0 0 4 0 】

以上より、絶対優先度判定部 3 2 j からイネーブル信号（有効信号）が発生すると、該イネーブル信号が入力したラウンドロビン回路 3 3 k i は、自身の待ち行列の先頭行列構成要素からリクエスト信号(Request)が入力されているか調べ、入力されていれば対応するキューイングバッファに対してパケットの転送許可(Grant)を出力する。又、設定されている伝送保証データ量のパケットデータの転送が完了し、あるいは、全待ちパケットの送出が完了し、連続転送許可部 3 1 d から巡回信号(Rotate)が入力すれば、先頭の行列要素を末尾に移動させ、残り

の行列構成要素を1つつ前進させ、これにより、次の行列構成要素を先頭に位置させる。

#### 【 0 0 4 1 】

##### (c) 全体の動作

以上より、図3，図4には8ポート ( $P_0 \sim P_7$ ) のスケジューリング部が示され、サービスクラス単位に待ち行列要素Queue-iGS, Queue-iCLS, Queue-iBE, Queue-iSpareが構成され、各待ち行列構成要素に応じたキューイングバッファQBFi<sub>1</sub> ~ QBFi<sub>4</sub>が設けられている。スケジューリング部24はこの待ち行列要素に対して帯域制御を行う。

各待ち行列要素に応じた送出待ちパッケージがキューイングバッファに存在するとき、リクエスト信号発生部31aは転送要求のリクエスト信号(Request)を発生してスケジューラ32に inputs する。絶対優先度判定部32jは所属する行列構成要素のリクエスト信号(Request)が少なくとも1つ入力されているグループのうち優先度が最大のグループをパッケージ送出グループとして決定し、該グループにenable信号(有効信号)を出力する。

#### 【 0 0 4 2 】

各優先度グループには、個別のRound-Robin回路33k<sub>1</sub> ~ 33k<sub>4</sub>が用意されておりenable信号が与えられたグループのRound-Robin回路33k<sub>i</sub>が有効となる。このRound-Robin回路33k<sub>i</sub>は、リクエスト信号(Request)を出力している行列構成要素の中から最も優先度の高い行列構成要素 ( $P_0$ -GSとする)を求め、該行列構成要素  $P_0$ -GSに対して転送許可信号(Grant)を出力する。転送許可信号(Grant)は、行列構成要素  $P_0$ -GSに応じた入力ポート  $P_0$ の該当するキューイングバッファQBF<sub>01</sub>にGrant回路31cを介して伝えられる。これにより、キューイングバッファQBF<sub>01</sub>は記憶するパッケージの転送を開始する。

この時、行列構成要素  $P_0$ -GSに対して予め設定されている伝送保証帯域値(伝送保証データ量byte/sec)TH<sub>1</sub>をレジスタ31fにロードする。これは、行列構成要素  $P_0$ -GSに対して一回の転送許可において許されている物理帯域を守るためである。基本的に行列構成要素  $P_0$ -GSに応じたキューイングバッファQBF<sub>01</sub>からのパッケージ読み出しは、実際の送出データ量が伝送保証データ量TH<sub>1</sub>と等しくなる

まで連続して行われる。すなわち、比較部 3 1 i により実際の送出データ量が保証データ量に達したことが検出されると、連続Grant回路 3 1 d は巡回信号(Rotate)をスケジューラ部 3 2 へ出力する。又、連続Grant回路 3 1 d は、実際の送出データ量が伝送保証データ量に達しなくても、待ちパケットをすべて送出してバッファが空になったことを検出すると転送許可状態(Grant状態)を放棄し、巡回信号(Rotate)をスケジューラ部 3 2 へ出力する。

#### 【 0 0 4 3 】

キューイングバッファQBF<sub>01</sub>から読み出されたパケットはフロー識別部 3 2 c により入力ポート番号/サービスクラスが識別され、行列構成要素はP<sub>0</sub>-GSのパケットフローであると識別される。これにより、スケジューラ部 3 2 における行列構成要素P<sub>0</sub>-GSに応じたカウンタ 3 2 b<sub>01</sub>はパケットの流量をカウントする。各カウンタ 3 2 b<sub>01</sub>~3 2 b<sub>74</sub>は同様にパケットデータの流量をカウントし、設定時間内における対応する行列構成要素に応じたパケットの流量を計測し、該設定時間がくれば自動的にクリアされる。各行列構成要素には予め設定時間における最大帯域(最大送出データ量byte/sec)が設定されており、この時間内にカウンタ 3 2 b<sub>01</sub>~3 2 b<sub>74</sub>の計数値が設定値と等しくなると、比較部 3 2 e<sub>01</sub>~3 2 e<sub>74</sub>は送出禁止信号INHを出力し、リクエスト禁止部 3 1 bはこの送出禁止信号により対応する行列構成要素に応じたリクエスト信号(Request)をマスクする。この結果、いくら行列構成要素に応じた送出待ちパケットが存在していても、スケジューラ部 3 2 からの転送許可は出力されなくなる。

#### 【 0 0 4 4 】

本発明では優先度グループの選択論理を単純な絶対優先度方式としている。単純な絶対優先度方式で良い理由を説明する。ここでいう単純な絶対優先度方式とは、「あるスケジューリング処理において、その瞬間の最も優先度の高いグループが必ず選択される」というものである。つまり、高優先度グループにある程度パケットが存在している場合、転送許可(Grant)はその高優先度グループに属する行列構成要素間にて巡回し、それより低優先度グループへ転送許可が渡される事はない。従って、この絶対優先度方式だけでは、最悪のケースの場合、低優先度グループの行列構成要素にパケット送出権が与えられない可能性が出てくる。

しかし、本発明では、各行列構成要素単位に流量カウンタ  $32b_{01} \sim 32b_{74}$  が用意されており、常に設定流量と実際の流量との比較処理を行っており、実際の流量が設定流量を超えた事が検出されると、その行列構成要素からの転送要求のリクエスト信号(Request)をマスクする。このため、設定流量に達した行列構成要素に対して以後設定時間内にスケジューラから転送許可(Grant)が出力されることはない。これにより、本発明によれば、各行列構成要素毎に実際の流量を監視して不当な帯域割り当てを防止できる。

#### 【0045】

本発明では、1回の転送許可による伝送保証データ量に実際の送出データ量が達しなくても、待ちパケットをすべて送出してバッファが空になったことを検出すると転送許可状態を解除し、他の行列構成要素に対して転送許可(パケット送出権)を渡すようにしている。このようにすれば、共有している物理帯域において無駄な空き帯域なくなり、帯域を100%利用する事が可能となる。しかし、この方式は、伝送保証帯域に達しないにも関わらず、転送許可状態を解除することを許した事になり、各行列構成要素に割り当てた帯域保証が守られるのかという疑問が生まれる。しかし、本発明では設定時間毎に前述のマスク制御により帯域を制御するため何らの問題は生じない。

#### 【0046】

##### (B) 第2実施例

図7は本発明の第2実施例であるn組の入出力ポートを有するパケットスイッチ装置の全体構成図である。このパケットスイッチ装置は、パケットを所定の伝送路に送出する際、ポート部毎に該パケットが到来した入力ポートとパケットのサービスクラスとに基づいて所定のスケジューリングを行って伝送路に送出する。

図7において、 $51_0 \sim 51_n$ は1組の入出力ポートに対応して設けられたポート部、52は各入力ポートから到来するパケットを所定の出力ポートにスイッチングするポート接続部(スイッチ部)であり、各ポート部 $51_0 \sim 51_n$ は同一の構成を備えている。

各ポート部 $51_0 \sim 51_n$ は、パケット入力方向に関して、入力側IF部(図示



せず) から入力するパケットに所定の処理を施して出力するフォワーディング部 6 1、フォワーディング部から入力するパケットを待ち行列(キュー)に接続して順番に出力する ingress 側キューイング部 6 2 を備え、出力方向に関して egress 側キューイング部 6 3、所定のスケジューリングに従ってパケットを出力側 I F 部から伝送路に送出するスケジューリング部 6 4 を備えている。

#### 【 0 0 4 7 】

フォワーディング部 6 1 は第 1 実施例のフォワーディング部  $22_0 \sim 22_n$  と同様に IP パケットに装置内ヘッダ FIH (Frame Information Header) を付加して出力する。

ingress 側キューイング部 6 2 はフォワーディング部から入力するパケットを待ち行列(キュー)に接続して順番に出力するもので、例えば 1 個のキューイングバッファを備え FIFO 形式でパケットを蓄積/出力するもの、あるいは、QoS のサービスクラス毎にキューイングバッファを備え、各バッファにサービスクラス対応のパケットを蓄積し、優先度の高いサービスクラスのパケットから出力するもの等がある。

スイッチ部 5 2 は  $n$  入力、 $n$  出力のスイッチであり、各ポート部  $51_0 \sim 51_7$  ( $n=7$ ) から入力するパケットをスイッチングして所定のポート部に出力する。

egress 側キューイング部 6 3 は多数のキューイングバッファを備え、スイッチ部 5 2 から出力するパケットを該パケットの入力ポート毎に仕分けると共に QoS のサービスクラス毎に仕分けしてそれぞれを各キューイングバッファに記憶する。入力ポート数を 8、QoS のサービスクラス数を 4 (iGS, iCLS, iBE, iSpare) とすれば、egress 側キューイング部 6 3 は各入力ポート毎に 4 個ずつ総計 3 2 個のキューイングバッファ  $QBF_{01}$ 、 $QBF_{02}$ , ...,  $QBF_{74}$  を備え、それぞれにパケットを記憶する。

#### 【 0 0 4 8 】

スケジューリング部 6 4 は egress 側キューイング部 6 3 の各バッファ  $QBF_{01} \sim QBF_{74}$  にキューイングされているパケットを所定のスケジュールに従って送出する。すなわち、スケジューリング部 6 4 は、各バッファに蓄積されたパケットに対して、帯域を管理しながらパケット単位に転送許可 (grant) を出力し、パケッ

トを出力側IF部を介して伝送路に送出する。スケジューリング部64は第1実施例と同様に図3、図4に示す構成を備えている。ただし、キューイングバッファQBF<sub>01</sub>～QBF<sub>74</sub>を含めて示している。

ここで注意すべき点は、入力ポートP<sub>0</sub>から入力したパケットはサービスクラス (iGS, iCLS, iBE, iSpare)毎に仕分けされてキューイングバッファQBF<sub>01</sub>～QBF<sub>04</sub>にキューイングされ、入力ポートP<sub>1</sub>から入力したパケットはサービスクラス (iGS, iCLS, iBE, iSpare)毎に仕分けされてキューイングバッファQBF<sub>11</sub>～QBF<sub>14</sub>にキューイングされ、以下同様に入力ポートP<sub>7</sub>から入力したパケットはサービスクラス (iGS, iCLS, iBE, iSpare)毎に仕分けされてキューイングバッファQBF<sub>71</sub>～QBF<sub>74</sub>にキューイングされる点である。スケジューリング部64の動作は第1実施例と同じである。

以上では、本発明をパケットスイッチング装置に適用した場合について説明したが、パケットルーチング装置にも本発明を適用することができる。

#### 【0049】

##### ・付記

(付記1) 優先度がそれぞれ異なる複数のグループに属するパケットを順に送出するパケット送出装置において、

グループ毎に待ち行列を生成し、各待ち行列を構成する要素に順番にパケット送出権を与える待ち行列制御部、

少なくとも1つの待ち行列構成要素に応じたパケットが送出待ちになっているグループのうち、優先度が最大のグループをパケット送出グループとして決定するパケット送出グループ決定部、

を備え、待ち行列制御部は前記決定されたパケット送出グループの待ち行列においてパケット送出権のある行列構成要素に応じたパケットを送出することを特徴とするパケット送出装置。

#### 【0050】

(付記2) 前記待ち行列制御部は、パケット送出権のある行列構成要素に応じたパケットが送出待ちになっていなければ次の順位の行列構成要素に応じた送出待ちパケットを送出する、

ことを特徴とする付記 1 記載のパケット送出装置。

(付記 3) 入力ポートと該ポートより入力するパケットに付加する各品質クラスとの組み合わせを待ち行列構成要素とし、それぞれの待ち行列構成要素が属するグループを設定するグループ設定部、

を備え、前記待ち行列制御部は、

待ち行列毎に各待ち行列構成要素にラウンドロビン形式で順番に、かつ、均等にパケット送出権を与える、

ことを特徴とする付記 1 記載のパケット送出装置。

【 0 0 5 1 】

(付記 4) 待ち行列構成要素毎に送出待ちのパケットを記憶するバッファ部、送出待ちパケットが記憶されているバッファ部に応じた待ち行列構成要素毎に送出要求のリクエスト信号を発生するリクエスト発生部、

を備え、前記送出グループ決定部は、各グループに属する待ち行列構成要素からのリクエスト信号の有無に基づいて送出待ちパケットが存在するグループを識別し、これらグループのうち優先度が最大のグループをパケット送出グループとして決定する、

ことを特徴とする付記 1 記載のパケット送出装置。

【 0 0 5 2 】

(付記 5) 行列構成要素毎に 1 回のパケット送出権で保証するデータ送出量を設定する保証データ量設定部、

パケット送出権を与えられた行列構成要素に応じたパケットの実際の送出量を監視する監視部、

実際のデータ送出量が前記保証データ量と等しくなったとき、送出権を次の行列構成要素に渡すための制御信号を出力する制御信号発生部、

を備え、前記待ち行列制御部は、該制御信号に基づいてパケット送出権を次の行列構成要素に付与する、

ことを特徴とする付記 1 または 3 記載のパケット送出装置。

(付記 6) 前記制御信号発生部は、実際の帯域が保証帯域を越える前にパケットがバッファからすべて送出されたことを検出したとき、前記制御信号を出力す

る、

ことを特徴とする付記 5 記載のパケット送出装置。

【 0 0 5 3 】

(付記 7) 待ち行列構成要素毎に設定時間当りに送出するパケットのデータ送出流量を設定するデータ送出流量設定部、

待ち行列構成要素毎に前記設定時間当りのパケットの実際のデータ送出流量を監視する手段、

待ち行列構成要素毎にパケットの実際のデータ送出流量を監視し、該データ送出流量が前記設定データ送出流量に等しくなった時、前記設定時間が経過するまで該待ち行列構成要素に応じたパケットの送出を禁止する送出禁止信号を発生するパケット送出禁止手段、

を備えたことを特徴とする付記 1 又は 3 記載のパケット送出装置。

【 0 0 5 4 】

(付記 8) パケット送出禁止手段は、前記設定時間毎に実際のデータ送出流量を零にクリアして送出禁止を解除すること、

を特徴とする付記 7 記載のパケット送出装置。

(付記 9) 前記リクエスト発生部は、所定の待ち行列構成要素に対して送出禁止信号が発生したとき、該待ち行列構成要素に応じた送出待ちパケットが存在してもリクエスト信号を発生しない、

ことを特徴とする付記 8 記載のパケット送出装置。

【 0 0 5 5 】

【発明の効果】

以上本発明によれば、ストリクトキューイング方式及び重み付けキューイング方式を併用して可変長パケットを分割することなく帯域制御を可能にし、スケジューラの回路規模の縮小化、高速処理化を図ることができる。

又、本発明によれば、入力ポートと該入力ポートより入力するパケットの品質クラスの組み合わせを待ち行列構成要素とし、パケット単位でなくこの待ち行列構成要素単位で帯域制御、優先度制御するようにしたから、スケジューラの回路規模の縮小化、高速処理化を図ることができ。

又、本発明によれば、パケット分割処理やパケット組立て処理を不要にすることができる。

【0056】

又、本発明によれば、1回のスケジューリング処理時間が短くなるように、すなわち、スケジューリング処理時間が最小パケットの伝送時間以内で収まるように構成できる。このため、たとえ最小パケットを連続して出力する状況においても、空き帯域なくパケット・データを連続して出力することができる。

又、本発明によれば、待ち行列構成要素を数段階の絶対優先度グループにグルーピングし、グループ内でRound-Robin方式により各待ち行列構成要素に順番にパケット送出権を与えるようにしたから、単純なRound-Robin方式であっても、優先度の高い特定の行列要素に対して、転送許可の巡回頻度を上げることが可能となる。

又、本発明によれば、優先度グループを単純な絶対優先度方式により選択することにより、簡単なハードウェアによる高速スケジューリング処理ができる。

【0057】

又、本発明によれば、各優先度グループに属する行列構成要素を任意に設定でき、また、その数に制限を設けないようにできる。しかも、優先度の高いグループの流量と、優先度の低いグループの流量の比率が偏っていても、帯域制御が可能である。

又、本発明によれば、行列構成要素毎に1回の転送許可による伝送保証帯域を設定するようにしたから、各行列構成要素に割り当てる帯域の制御及び帯域の管理が可能となる。

又、本発明によれば、送出する待ちパケットが存在しなくなれば、直ちにその行列構成要素に対する転送許可を解除し、他の行列構成要素に転送許可を渡すようにしたから、共有している物理帯域において無駄な空き帯域をなくすことができる。

【0058】

又、本発明によれば、各行列構成要素毎に所定時間当りに送出するパケットのデータ送出流量（帯域）を設定し、各行列構成要素毎にパケットのデータ送出流

量を監視し、該データ送出流量が設定データ送出流量を越えた時、所定時間が経過するまで行列構成要素に応じたパケットの送出を禁止するようにしたから、厳密な帯域制御を実現できる。

【図面の簡単な説明】

【図 1】

本発明のパケット送出装置の構成図である。

【図 2】

装置内ヘッダの構成図である。

【図 3】

スケジューリング部の詳細構成図（その 1）である。

【図 4】

スケジューリング部の詳細構成図（その 2）である。

【図 5】

優先度グループの例である。

【図 6】

ラウンドロビン方式の説明図である。

【図 7】

本発明のパケットスイッチ装置の全体構成図である。

【図 8】

単純優先度方式の説明図である。

【図 9】

単純優先度方式の別の説明図である。

【図 1 0】

重み付けキューイング方式の説明図である。

【図 1 1】

重み付けキューイング方式の別の説明図である。

【図 1 2】

重み付けキューイング方式の問題点説明図である。

【図 1 3】

重み付けキューイング方式の問題点解決手法の説明図である。

【図 1 4】

重み付けキューイング方式の解決手法の問題点説明図である。

【符号の説明】

1 0・・・パケット送出装置

1 1<sub>0</sub>～1 1<sub>n</sub>・・・入力伝送路

1 2・・・出力伝送路

2 1<sub>1</sub>～2 1<sub>n</sub>・・・入力側 I F 部

2 2<sub>1</sub>～2 2<sub>n</sub>・・・フォワーディング部

2 3<sub>0</sub>～2 3<sub>n</sub>・・・ingress側キューイング部

2 4・・・スケジューリング部

2 5・・・egress側キューイング部

2 6・・・出力IF部

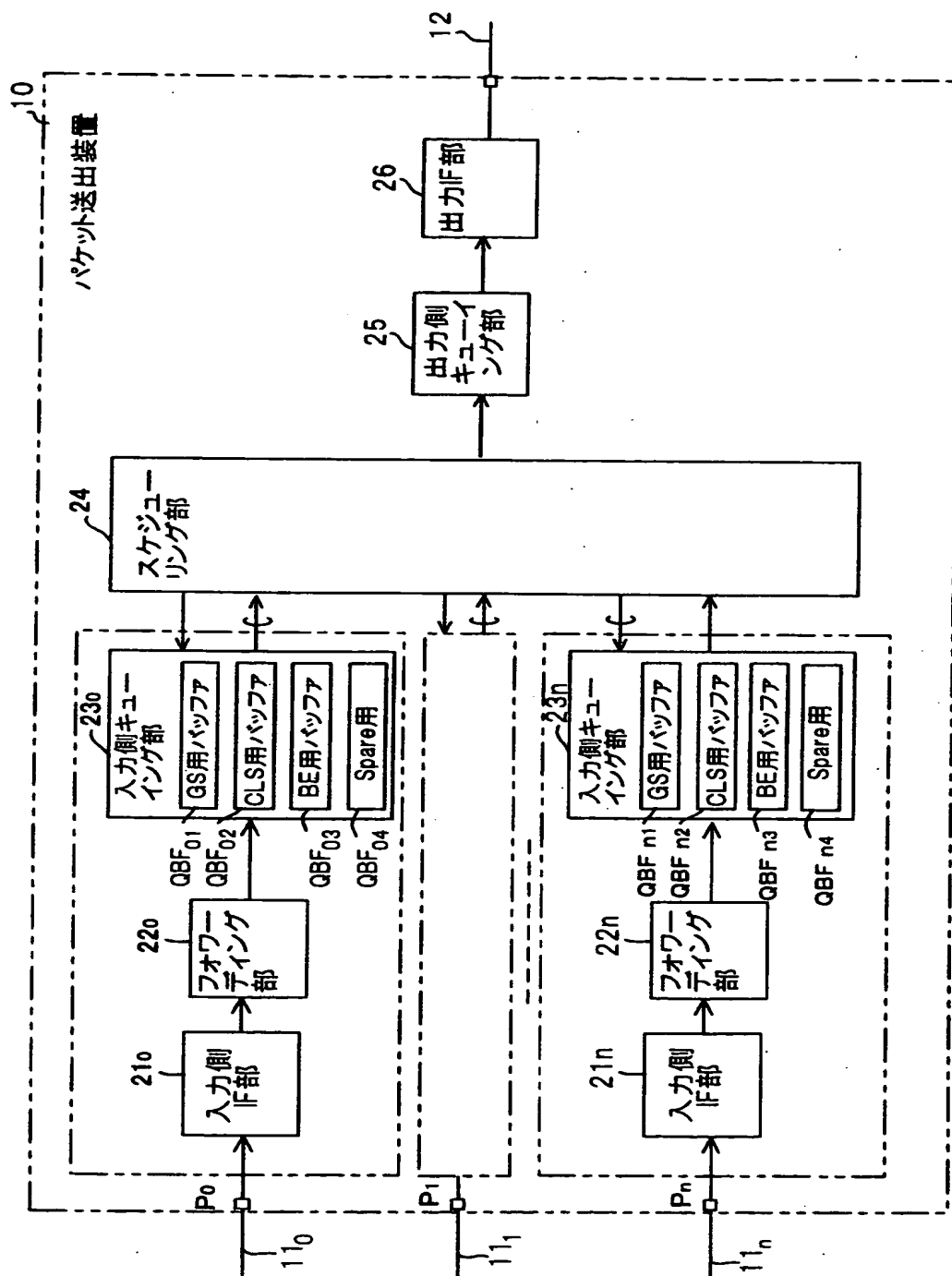
P<sub>0</sub>～P<sub>n</sub>・・・入力ポート

QBF<sub>i1</sub>～QBF<sub>i4</sub> (i=0～n)・・・キューイングバッファ

【書類名】 図面

【図1】

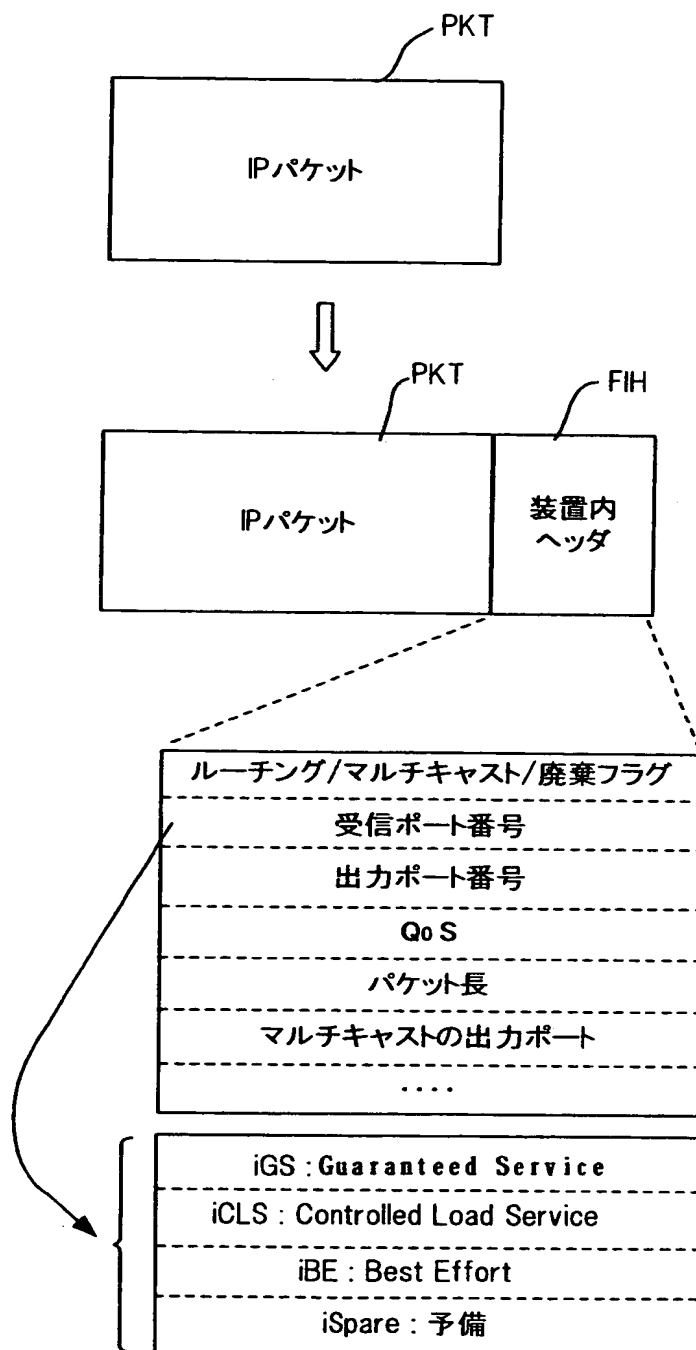
本発明の packets 送出装置の構成





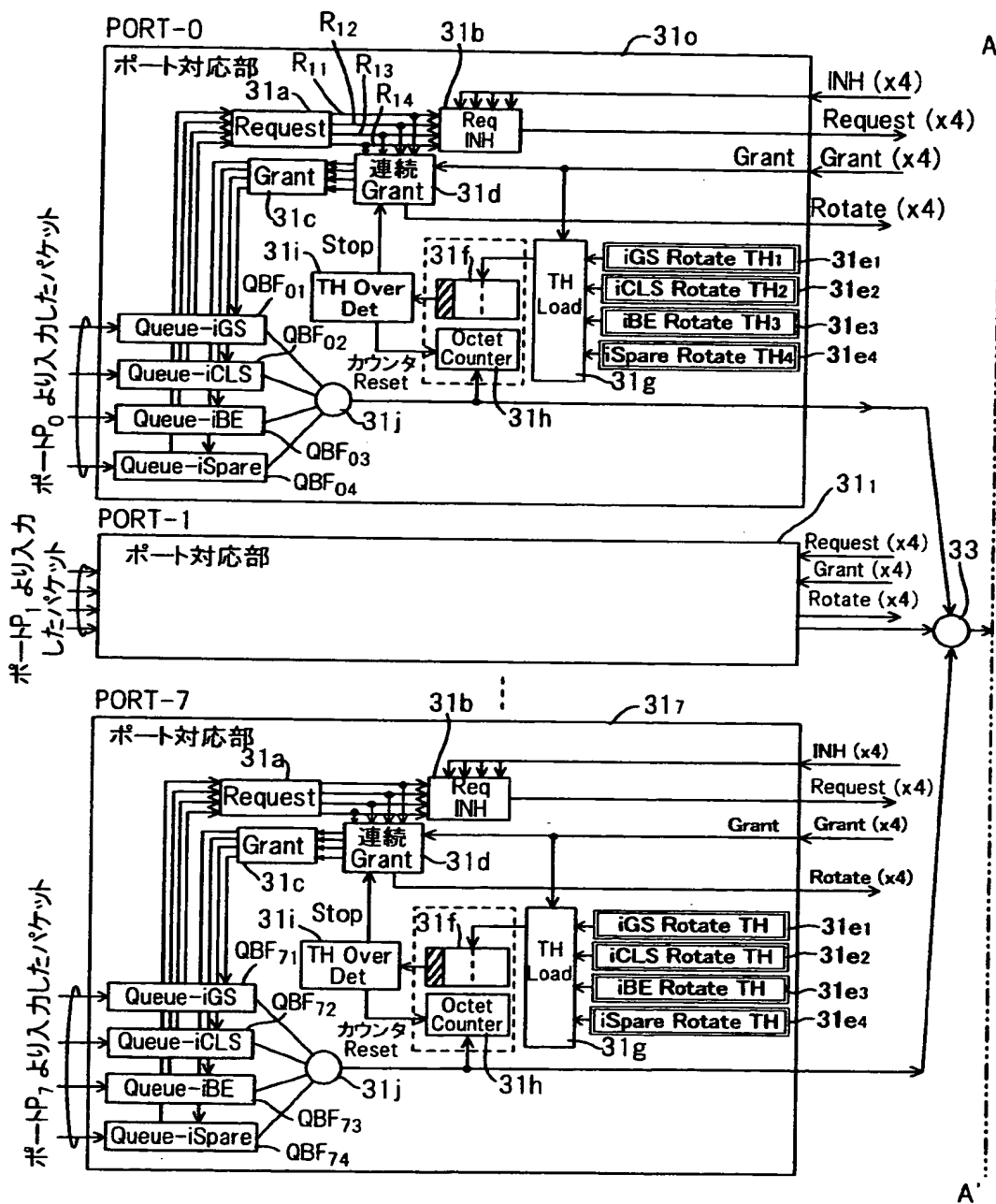
【図 2】

装置内ヘッダの構成



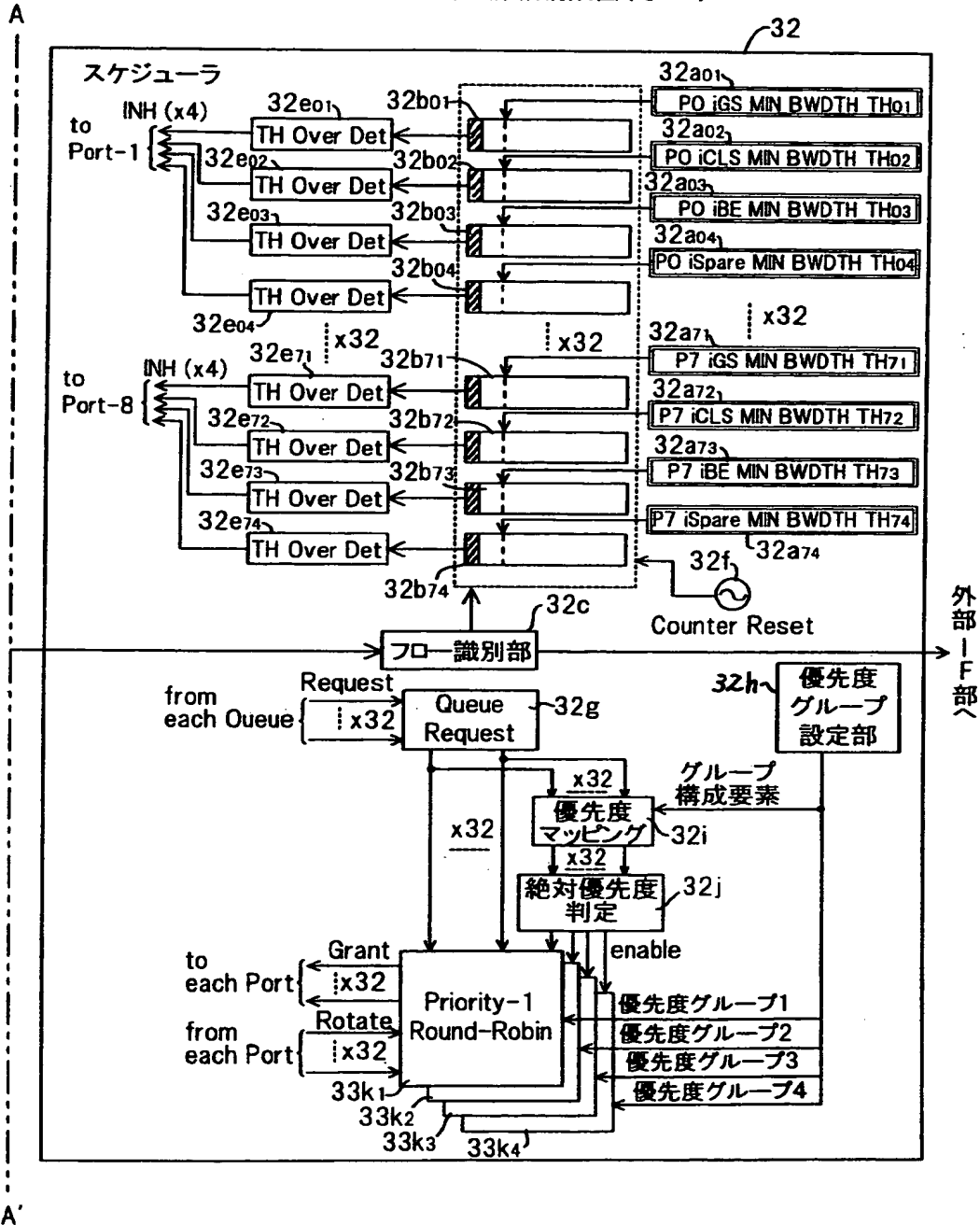
【図 3】

### スケジューリング部の詳細構成図(その1)



【図 4】

スケジューリング部の詳細構成図(その2)



【図 5】

## 優先度グループの例

(A)

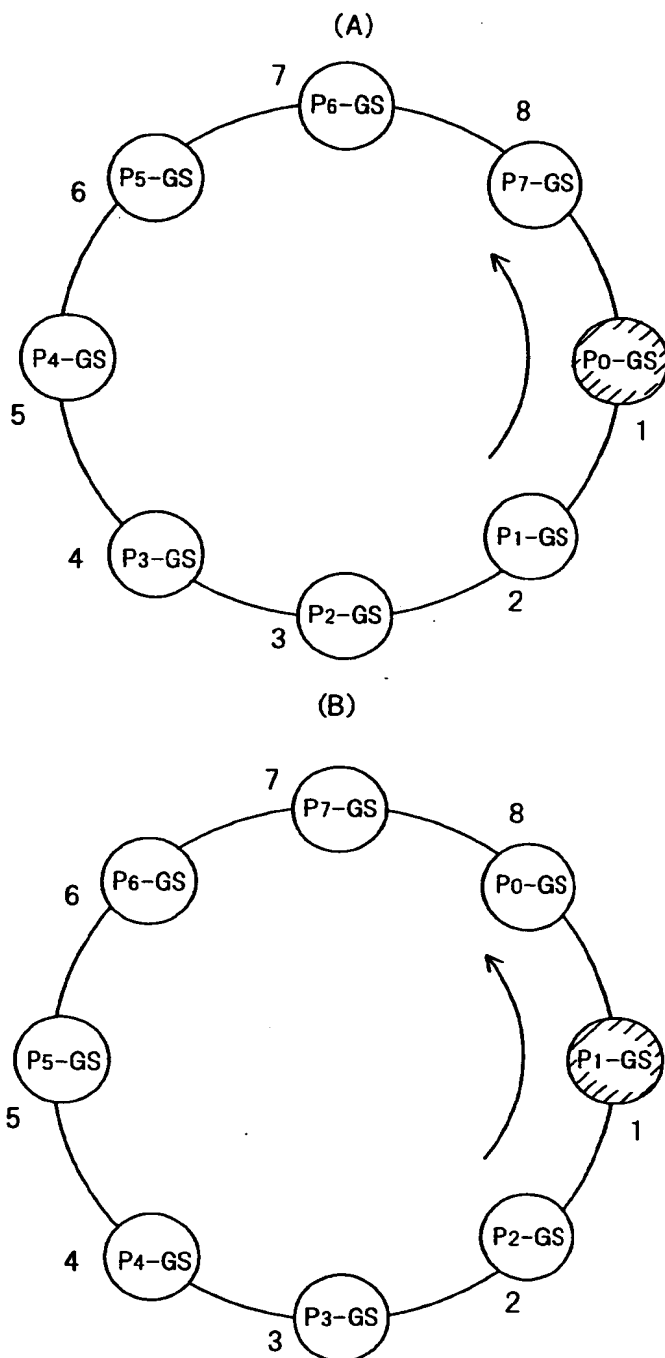
優先順位グループ	ポートNo. & QoS	個数
第1優先グループ	P0-GS ~ P7-GS	8
第2優先グループ	P0-CLS ~ P7-CLS	8
第3優先グループ	P0-BE ~ P7-BE	8
第4優先グループ	P0-Spare ~ P7-Spare	8

(B)

優先順位グループ	ポートNo. & QoS	個数
第1優先グループ	P0-GS ~ P6-GS	7
第2優先グループ	P7-GS P0-CLS ~ P7-CLS P0-BE	10
第3優先グループ	P1-BE ~ P7-BE	7
第4優先グループ	P0-Spare ~ P7-Spare	8

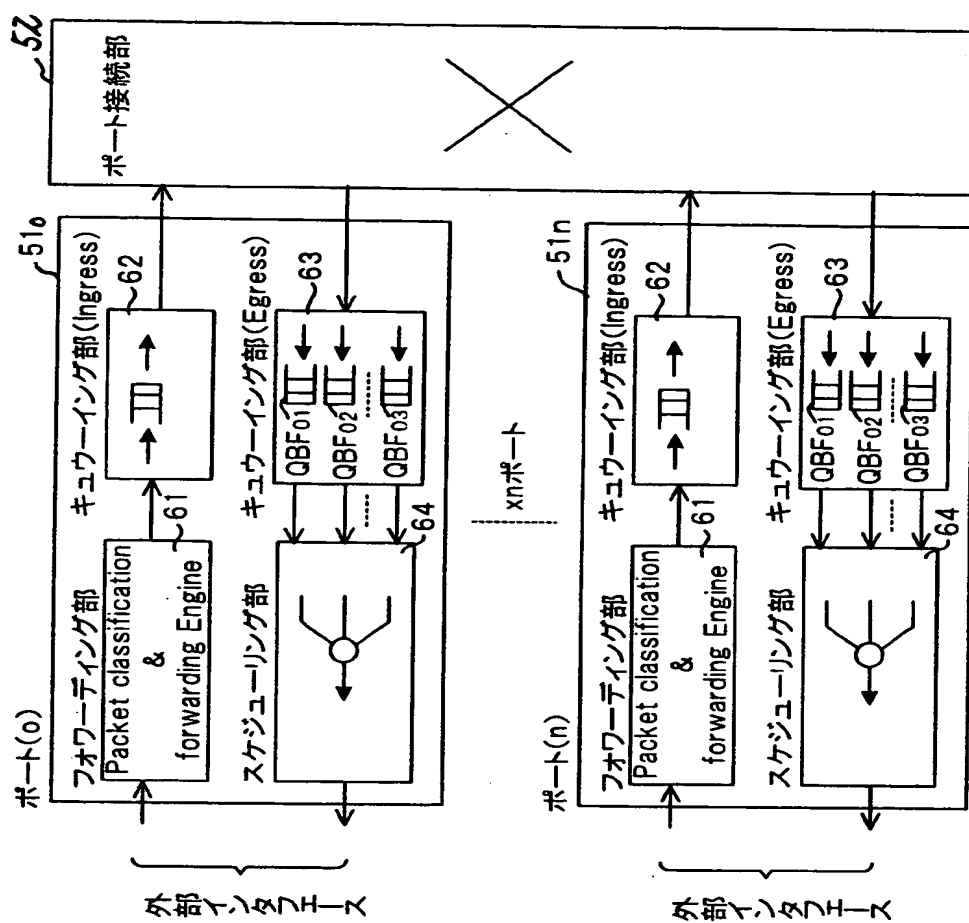
【図 6】

ラウンドロビン方式説明図



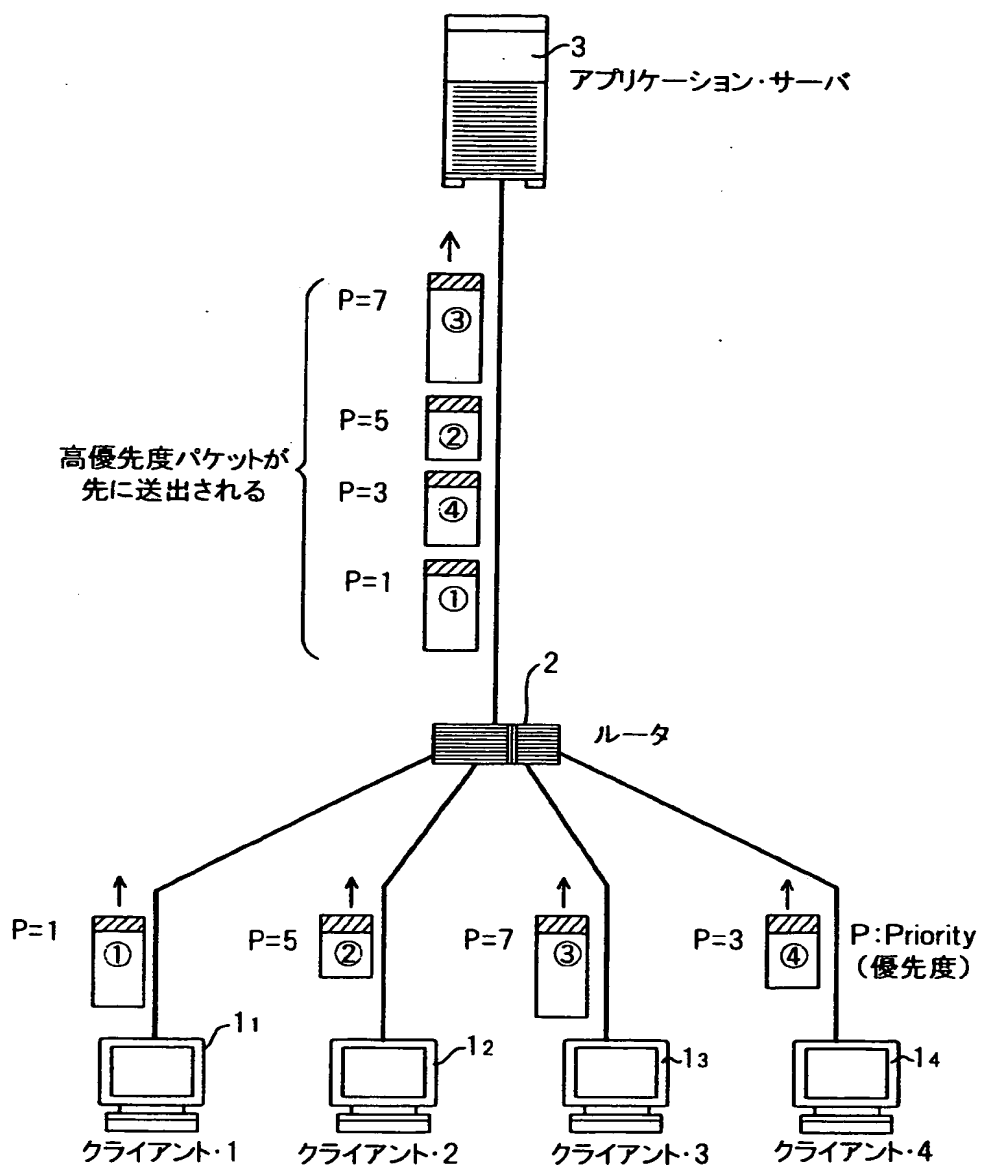
【圖 7】

## 本発明の PACKET SWITCH 装置の全体の構成



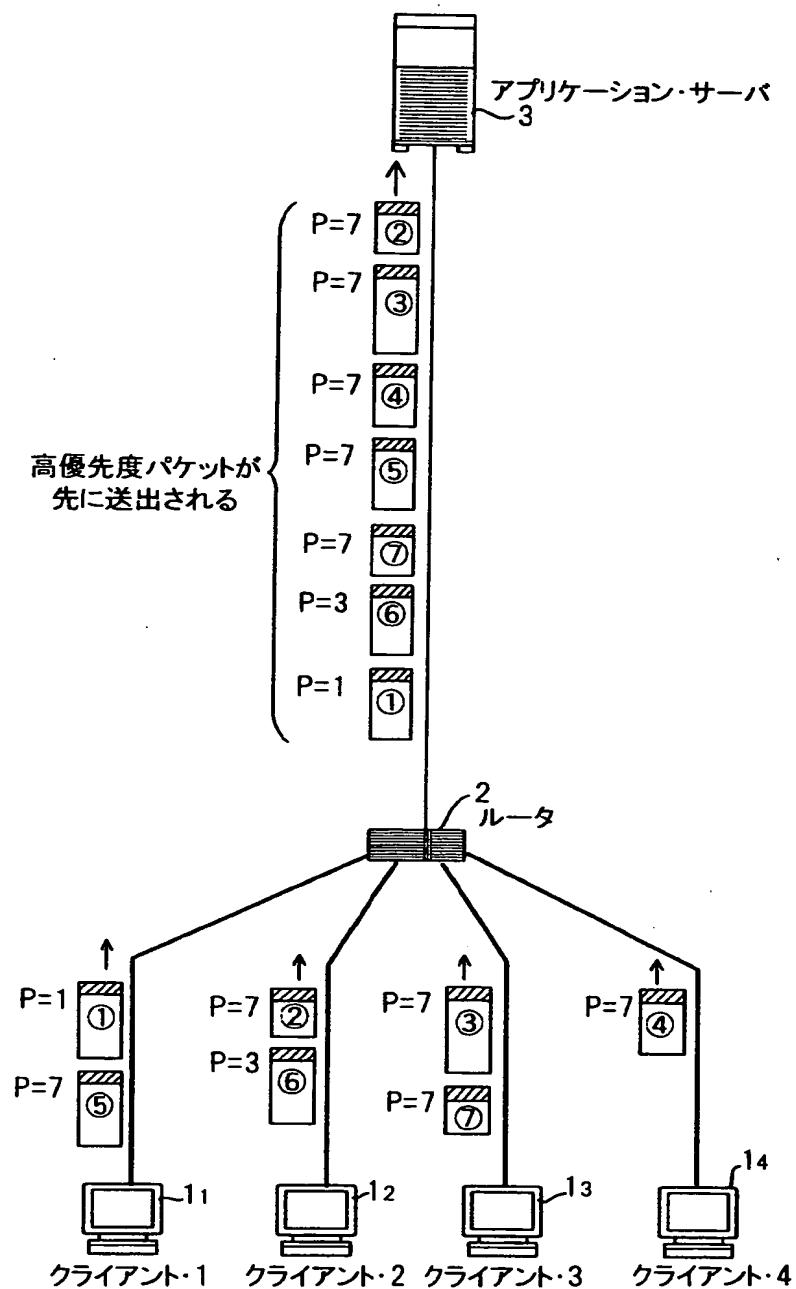
【図 8】

単純優先度方式説明図



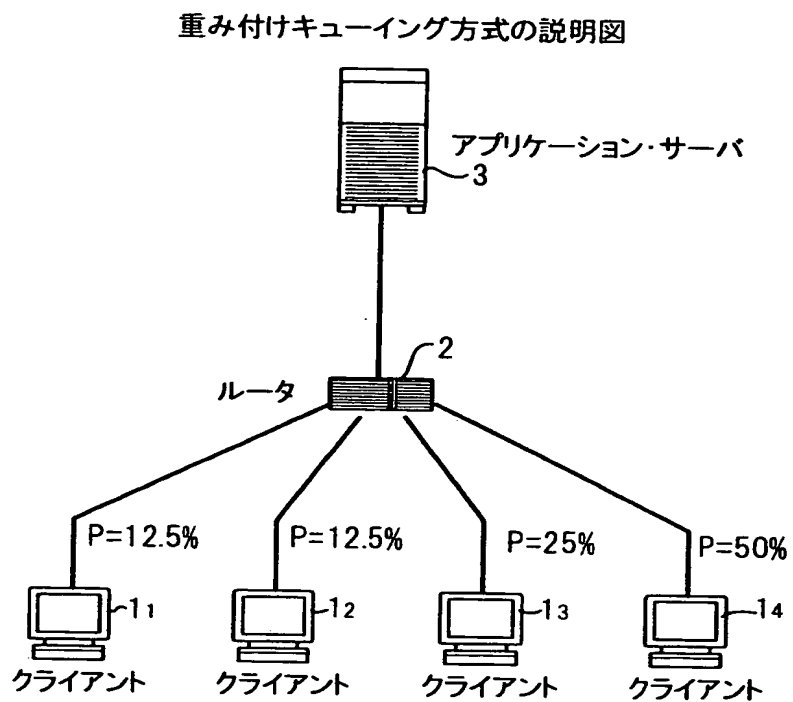
【図 9】

単純優先度方式の別の説明図

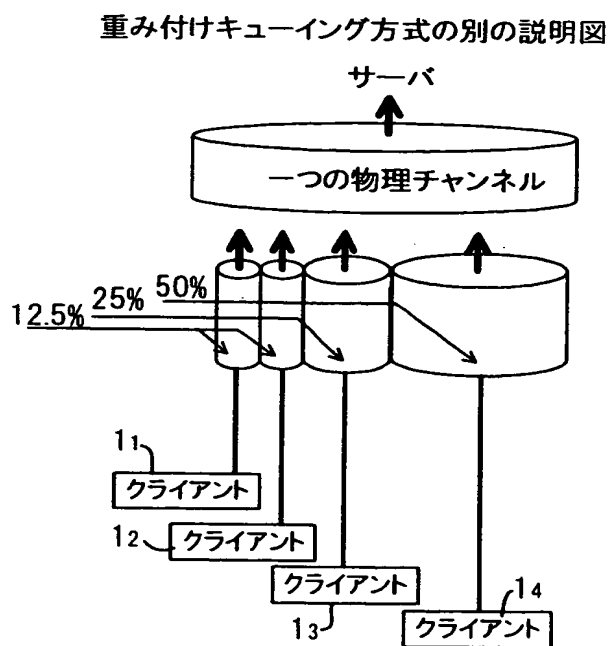




【図10】

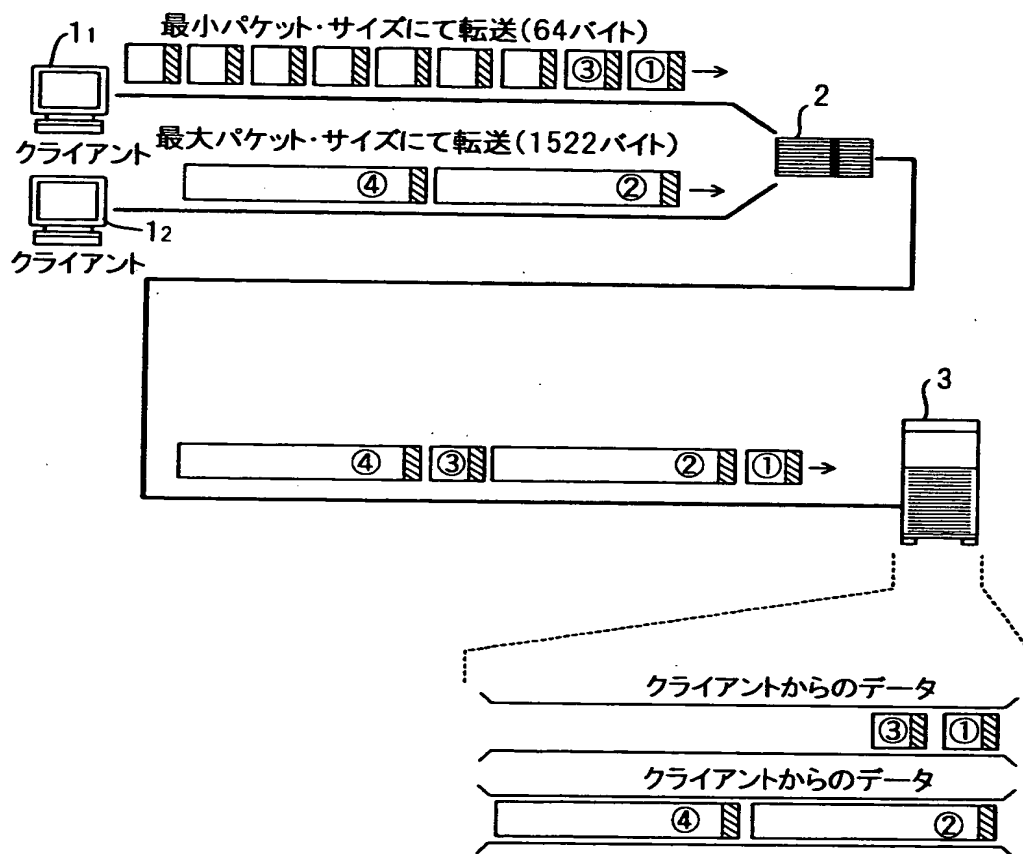


【図11】



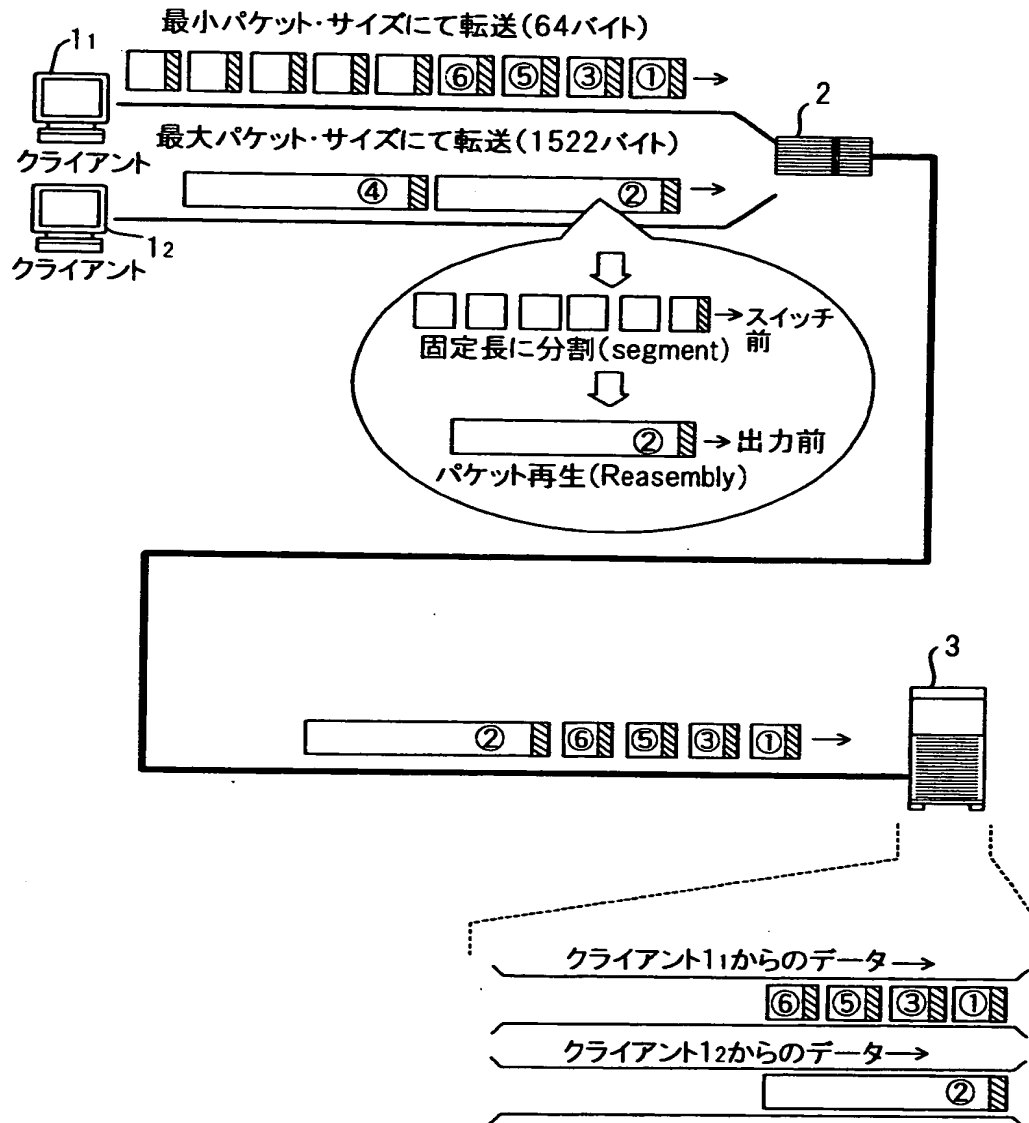
【図 12】

重み付けキューイング方式の問題点説明図



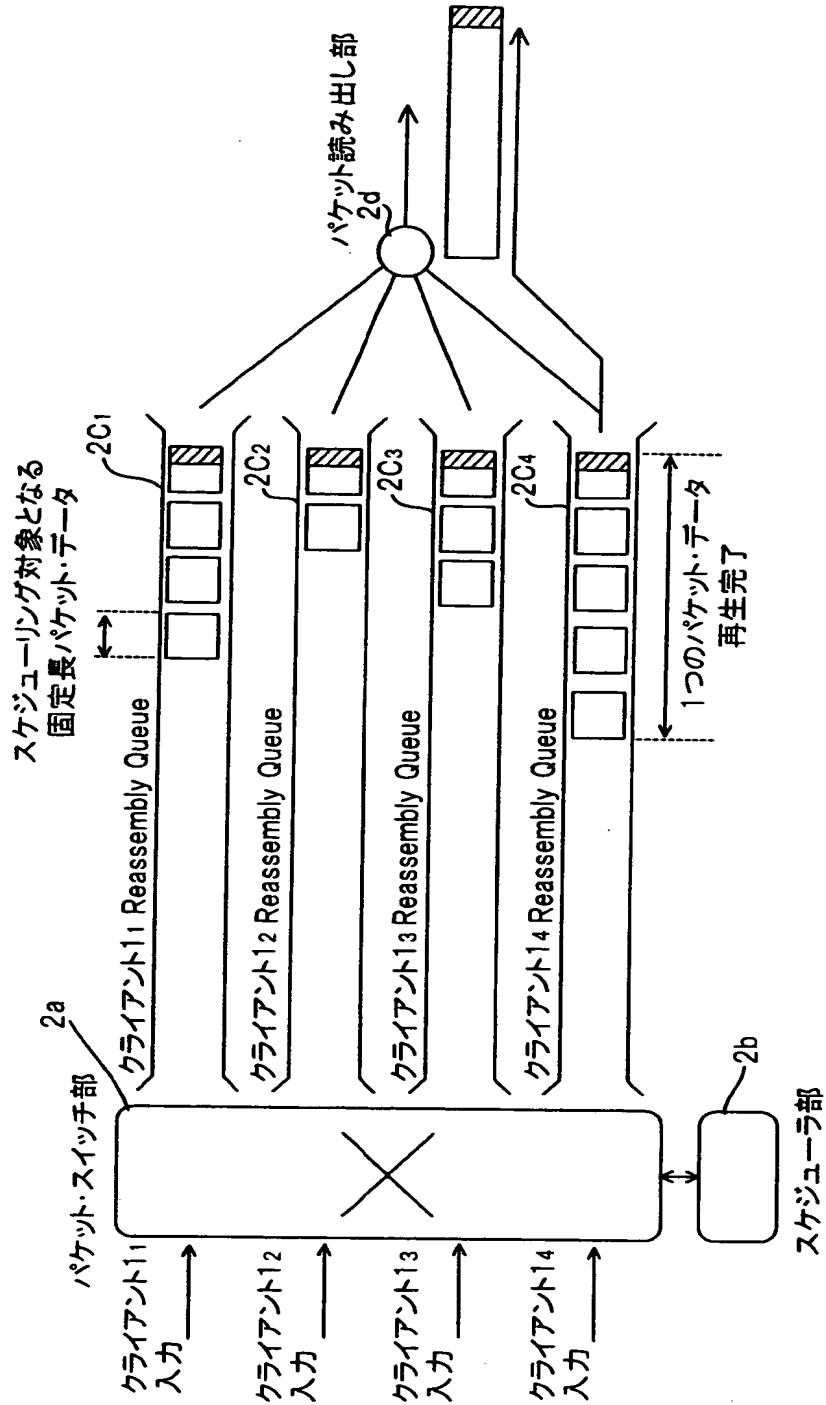
【図13】

重み付けキューイング方式の問題点解決手法の説明図



【図 14】

重み付けキューイング方式解決手法の問題点説明図



【書類名】 要約書

【要約】

【課題】 可変長パケットを分割することなく帯域制御を可能にする。

【解決手段】 優先度がそれぞれ異なる複数のグループに属するパケットを送出するパケット送出装置である。このパケット送出装置において、スケジューリング部 2 4 の待ち行列制御部はグループ毎に待ち行列を生成し、待ち行列毎に該待ち行列を構成する要素にラウンドロビン方式に従って順番にパケット送出権を与える。スケジューリング部 2 4 の送出グループ決定部は、少なくとも 1 つの待ち行列構成要素に応じたパケットが送出待ちになっているグループのうち、優先度が最大のグループをパケット送出グループとして決定し、待ち行列制御部は決定された送出グループの待ち行列の最優先行列構成要素にパケット送出権を与え、該行列構成要素に応じたパケットを送出する。

【選択図】 図 1

認定・付加情報

特許出願の番号	特願2001-015615
受付番号	50100094337
書類名	特許願
担当官	内山 晴美 7545
作成日	平成13年 2月 6日

<認定情報・付加情報>

【特許出願人】

【識別番号】

000005223

【住所又は居所】

神奈川県川崎市中原区上小田中4丁目1番1号

【氏名又は名称】

富士通株式会社

【代理人】

申請人

【識別番号】

100084711

【住所又は居所】

千葉県千葉市花見川区幕張本郷1丁目14番10

号 幸栄パレス202 齋藤特許事務所

【氏名又は名称】

齋藤 千幹

出 願 人 履 歴 情 報

識別番号

[000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号

氏 名 富士通株式会社